

Assessment of Large Student Cohorts In a Formal Distributed Learning Environment

**Naveed A. Malik
Founder-Rector
Virtual University of Pakistan**

It really is good to be back. I am going to talk about a real practical problem that we have faced, and how we have tried to use information technologies to tackle it. I will give you a brief background of the university, 30 seconds worth. We have already spoken about it many times. Then I will define the problem, and I would like you to appreciate the dimensions of the problem before we go into the solution, which is quite straightforward, but innovative. I'm going to talk about the fact that we are a virtual university. We use e-Learning online, television broadcasts and internet, for interaction. But what is critically important— and I mentioned this, I think in our 2005 talk— is that perception is actually defining reality.

What we need to do is to make sure that our degrees are accepted, and that happens if we have complete transparency. As far as our courses are concerned, they comprise video lectures. These are already being broadcast over free-to-air television, so in a way, we have opened the doors of our classroom, and said, “There's the professor, there's the content, look at the quality.” That is the first half.

The second half is, how do we assess the students? So we decided, as a design principle at the Virtual University from day zero that the assessments would be done in a very conventional manner. We will define exam days. It is a semester-system-based university, so we have midterm and final term examinations. Students would physically come into an exam center, sign up in a proctored environment, and sit down and take their test. It would always be conventional, visible and above board. Formal, proctored examinations would have to be done twice a semester, conducted at dedicated exam centers, face to face.

Let's look at the dimensions of the problem. When we started out with a 500-student cohort, it was fairly straightforward. Now, we are talking about 60,000 students, all over the country. We are in more than 95 cities. I can't even name some of them. We have an admission intake every semester. In spring and fall, we induct new students, which implies that all courses are being offered every semester. This is important to understand. In a normal, once-a-year induction, you have the spring courses and you have the fall courses. With twice-a-year induction, it means every course is being offered. We are talking about 160-plus courses.

There are additional dimensions to the problem. We have full-time students, young high school graduates who come to the university and are taking a full load -- five or six courses. They normally follow a defined schedule, which is fairly easy to handle. Then we have working professionals who have come in as students at the Virtual University.

These working professionals are taking a course load that suits them, sometimes a partial load. Sometimes they take one course, and then come back and take three or four in the next semester, and so on and so forth. So there is a variety of course distribution that happens. These professionals, of course, have limited study times and exam dates. Of course, in any university you find stragglers. In a virtual university, you find more, and they can take essentially any course in any semester.

Our problem basically began with the fact that as the student numbers increased, how could we actually stay with the academic calendar? In other words, how do you complete a semester on time, having to conduct two formal examinations? We have 60,000 students, each one taking three, four, five, six courses, which are happening all over the country, in 95 cities or more. We are trying to conduct these courses in a time-bound manner, without losing the sanctity of the examinations. Do you distribute paper? Do you do that on a computer? How do you do that? The major issue was the exam schedule. I think everybody who is into academics would understand that creating an exam schedule that satisfies all the students is almost an impossibility. Somebody is always going to complain, “Well, you've given me an exam in the morning, and one in the afternoon. Couldn't I have a break between papers? I would like to pace them out,” and things like that.

In our case, we have an issue in terms of the mix of student population that also brings an additional complexity into the picture. We had to worry about secure distribution of question papers across the country, maintaining the sanctity, efficient and secure collection of answer sheets, and grading within an acceptable time frame, because students need to know their results before they move on. We have a series of conflicting demands. We would like to run the examination tent, like this little auditorium, in full capacity. Every seat should be occupied. Therefore we can minimize the number of days it takes to conduct the exam. But working professionals coming in would say, “Look, I can't take every day off for the exams. I would rather do that in the evening.” Or, “Can you examine me on Sundays?” Full-time students— “10:00 a.m., Mondays, Wednesdays, and Thursdays. That's when we would like to sit for the exams.”

Then we have additional complexities— commuting students. In the interior part of Sindh, which is one of the southern provinces of the country, we've got students who have to commute about two hours to come to an exam center. They say it can't be very early. It can't be very late, because if girl students are traveling, “We don't want to travel at night, and we can't leave at dawn.” There are many dimensions to the problem. And of course, many student would like to pace out there examinations, so they would like to have gaps between papers.

From a management perspective, we have something more to look at. If I use the same question paper across the country, an exam in Islamabad starts a little earlier than that same exam in Karachi. Okay, this is the e-world; the questions have already been transmitted. The kids in Karachi, before they even enter the exam hall, have the question paper in their hands, so the sanctity is compromised. Ensuring proper invigilation all across the country is a nightmare. It is impossible, for example, for us to find IT-literate examiners or proctors. The kids are sharper on the computers. They would be running our

presentation in front, and chatting in the background: “Okay. This is the question. What is the answer?” Somebody in cyberspace is helping them along. We put in blocks on the servers, and things like that, but there was an interesting situation there as well. So, what is the idea?

The idea was that they were complaining too much about the exam schedule. We call it the date sheet. Why not let them create their own date sheets? So you decide, “I’m taking Programming on Monday morning at 9:00 a.m.” Somebody else says, “I’m doing Economics on Monday morning. I’ll take Programming on Tuesday evening, at 6:00 p.m.” So how do you manage that? The first thing is, if you allow the students the flexibility of creating their own exam schedule, first of all, the noise of the complaints goes away. Everybody is choosing a time and place of their own convenience.

But there are some requirements. The other dimension that is written here is, “first come, first served.” This makes sense. You find a seat of your convenience. If you don’t find a seat, you were too late. Then you sit wherever it is possible for you sit. The other one is to try and reduce the time taken to conduct the exams. We needed to run the exam centers at full capacity. Now, in a conventional system, when we say, “All right, Economics is Monday morning,” and three students are taking economics -- so a full hall has just three students sitting for the exam. Let’s say an introductory course in computers might have a full hall. You have this fluctuating attendance in the exam centers, and this actually is increasing the number of days or hours required to conduct the exam -- when we would like to run it totally, fully occupied. So what do we do?

If we allow the students to create their own individual date sheets, on a first-come, first-served basis, it can only work if the students receive distinct question papers. Because if you think about it, it implies that every potential exam session has every subject being examined. So at 9:00 a.m. every Monday morning, you’ve got economics, Islamic studies, Pakistan studies, Mathematics and Computer Science being examined. At 11:00, the same thing repeats. Tuesday morning, the same thing repeats. So we have to have distinct question papers. How do you create distinct question papers? You use a question bank. So we started, about two and a half years ago, defining, creating, and then populating a question bank, and then automating the process. So in terms of key concepts of what we have done, we have created a question bank.

We have students create their own individual date sheets. We have distinct paper generation, so each student gets a different question paper. We have all kinds of encryption and security measures in place, because we need to distribute it to 95 different cities or more, and then bring them back. We have to conduct the exams, and we have to grade them. Anybody who has graded exam papers would understand that there is a hidden nightmare sitting here, in that every question paper or answer script that you are trying to grade is different— I mean, the brain goes crazy when you try to do that. Then we have to declare the results in a time-bound, efficient manner. Let’s take a look at the question bank. It includes multiple choice as well as essay-type questions. We have categorized them and identified them according to which topic of the syllabus they cover, which lecture they cover, the difficulty level, and which cognitive skills they challenge. We have also determined how much time it would take to solve, how many points would

be awarded for the answer, and then a grading rubric to go along with the question. All of that goes through a quality-assurance process, and then the question is inducted into the question bank.

Throughout the semester, we have new questions being added daily. Every tutor, every instructor, has to contribute three to four questions every week into the question bank, so the bank is constantly being replenished and renewed. Each question carries a use count. Since the paper-generation process may use a certain question ten different times during an exam cycle, the question would have a use count-up. Beyond a certain point, we would retire that question, saying, "The entire student community knows about it. Let's take it out of circulation."

Planning the examination is an interesting process. We have to determine the paper slots. The students, times the number of subjects each student is taking, is the number of seats that we require in the exam sessions across the board. Everything is conducted on a computer. It is not a paper and pencil exam, but it does include essay type and all kinds of different questions. The capacity of the exam center is really the number of computers available, and we have a five-percent buffer for breakdowns and things like that.

Therefore, we can straight off calculate that we've got a capacity to examine 60,000 students in 12 days. That is the exam duration. Then we say, from morning to evening, we will have a session at 9:00 a.m., another one at midday, another one at 3:00 p.m., and one for working professionals at 6:00 p.m. We can have four sessions, "n" days. We can actually compute exactly what our capacity is to conduct an examination.

In terms of making a date sheet, students log in to a particular website, which is going through the learning management system itself. They use their own credentials: log-in name and password. Interestingly, on the first line, they choose the city where they want to be examined. For example, they may be studying in Karachi, but it turns out that they are attending a wedding of a relative in Lahore, and they will be in Lahore during that period. So they can get examined in Lahore, no problem. It is a big thing for the students when we let them choose the city in which they are going to be examined, but practically speaking, we find only a tiny percentage actually change cities. They normally get examined where they are studying.

Within the city, we have many exam centers. So Karachi and Lahore have multiple exam centers, and the students will have the facility to choose that center. Obviously, it is an IT-based system, so the system will display exactly the subjects that they are registered for during the semester. They would start defining, on day x, y, z— "On that date, at this time, I would like to sit for that exam." And, of course, the system is keeping a tally of the attendance in that center, so it can type easily and tell them, "Sorry, this session is full, choose another session." So it is a first-come, first-served making of a date sheet. For each subject, they can select a day and a session and then confirm. So it is entirely in their own hands, and they can't come back to us and say, "Look, you're giving me exams that are too close together."

So here is a view of the interface. There are some instructions, and you go and log in. To protect privacy, I have blanked out this student's face and his ID, but if you look, you will see that he did it right: He took Wednesday, Saturday, Wednesday, Saturday, Wednesday, Saturday— six courses, very well paced out. He selected 10:30 a.m. for all of his papers, a very convenient time. He was one of the early birds—came in, registered, got his date sheet set. Last person aboard takes the 5:00 a.m. bus.

Once this date sheet has been defined, we have all the information we need. We know that in Karachi on May 26th -- which is the next midterm examination - at 9:00 a.m., I need ten Introduction to Programming question papers, I need 15 Economics 101, I need 25 English Comprehension. I have complete statistics at my disposal. We do this process of date sheet generation about ten days before the start of the exam. The students have come, created their schedules and gone away. We get busy. The academic staff will provide the question paper parameters.

Maybe I can go into the details later, but the parameters are more like, "During this midterm exam, we would like to cover lectures one through 22, and within these lectures, I would want to have about 30% multiple choice questions, five short-essay, two medium-essay, one long-essay question." Once the teacher defines those parameters, the system already knows what the total score on the exam is, what the total duration of the exam is, and that's it. It does not matter if the final exam is worth 62 points. It does not have to be 100 points. It does not matter if it is 93 minutes. It is 93 minutes across the board. We do not really worry about those details, but it all works out. Then the system gets to work, and starts churning out these distinct question papers. There is a complex algorithm at the back that makes sure that there is an even distribution across the syllabus of every selection that is made. This generates 250,000 or 350,000 different question papers as required.

This entire process, by the way, is happening in an encrypted fashion. All the questions, all the data that's sitting inside the database, is also encrypted. So you can't read it. Nobody can read the question. The question just has an ID, and some gibberish in front of it. All e-papers then for a particular exam center are collected. You create an electronic bundle that is sent out by email, with an encrypted attachment. There are multiple layers of encryption in there. It is actually a database back-up file that goes out. It is burnt to a CD, and physically dispatched by courier -- just in case somebody has problems receiving a 10-megabyte attachment. There are sometimes issues in terms of connectivity of bandwidth, so CDs are sent out.

What happens at the recipient end -- the exam center end -- is that the center superintendent loads the CD on the exam server. They can load it; they can't open it. It can only be opened on the first day of the exams. It is opened through a password, and the password delivery mechanism is separate. It happens on the day of the exam, either by email to the superintendent or via SMS text message to the superintendent, each one getting a different password for that session. You've got lots of levels of security.

Once the file has been loaded, then only an individual session can be unlocked. Monday morning at 9:00 a.m. You cannot unlock the noon session because only the 9:00 session

will happen. Then as I said, the questions are being encrypted anyway. The other interesting part is, if on Monday morning at 9:00 a.m., we had ten kids in one particular city, in one particular examination center, taking the Introduction to Programming exam, there would obviously be ten question papers on the database server. But they still do not know, and we do not know, who is going to get which paper. It is only when they actually log into the examination application that one out of those ten is randomly taken and delivered to the student. It is stamped inside with the student's ID and is associated with the student, so it is quite random.

At the end of each day, rather than at the end of each session, we collect all of the student attempts. Again it is all electronic. We are talking about technologically challenged proctors, and superintendents, so we give them one big button on the screen, "Press here." It collects all the data, encrypts it, puts it into a bundle, attaches it to an email, and sends it out. The superintendents do not have to do anything. They just press a button.

Upon receipt, what do we do? This comes back to the academic grading question. We have an army of tutors and instructors waiting to grade. They are dreading the fact that every question paper is going to be different. "How do I get my brain to function?" What do we do? We take all the answer attempts from the students and we separate them into piles of identical questions. Since the questions were created from a question bank, there is a certain amount of reuse that is happening. So if question ID 1023 was used on Monday morning and Tuesday afternoon and Thursday evening, all the question numbers 1023 would be put into a single pile. And that one pile is given to a grader. You are grading the proof of the binomial theorem. Start in the morning, and until evening, you are just grading the proof of the binomial theorem question. So you get a lot of consistency in the grading.

It's not just that. As soon as the grader brings up a question to grade, remember the grading rubric that went into the question bank at the beginning? That comes alongside. So the first indication is, if the student did this right, give them two marks. If they missed this step, subtract a mark. So in terms of consistency of grading, it is very simple. You do not have to worry about switching gears, talking about a program at one, and a theory at the other, et cetera. You grade one question. When you are done, you will be given another question to grade.

It also leads to complete anonymity. The graders do not know whose attempt they are grading. They just have a proof, or a derivation, or a code or an essay in front of them.

Of course, IT again takes over. The result is collated by the system. Magically, when you say "Go," it appears in the student's grade book. The whole thing is done using IT. Semester- and subject-wide positions and campus-wide results are declared. And it worked. IT works.